

The Washback Effects of English Listening Test in Korean CSAT on Secondary School English Learning and Teaching*

Haeng Jung
(Honam University)

Jung, Haeng. 2008. The Washback Effects of English Listening Test in Korean CSAT on Secondary School English Learning and Teaching. *The Linguistic Association of Korea Journal*, 16(3), 143-162. This study investigated washback effects of English listening test in Korean CSAT on English teaching and learning in Korean EFL secondary schools. The following research questions were raised for this study. 1) What are the CSAT English listening test's characteristics? 2) What are the teachers' perceptions about the CSAT listening test and its effects on English teaching and learning? 3) What are the influences of the CSAT listening test on the English teaching and learning practices in secondary English classrooms? To answer the research question 1, the test items conducted over 13 years from 1994 to 2006 were collected and analyzed in terms of 5 categories such as language functions, topic, situation, text type, and factual/inferential understanding. For the research question 2, questionnaires were administered to 385 high school teachers and interviews were conducted to 38 high school teachers. For the research question 3, the data were obtained through a total of 8 class observations. The results earned from the analysis of the collected data indicated that the CSAT English listening test had effects on teaching contents, activities and materials in secondary classrooms.

Key Words: Testing, English Language Testing, Listening Test, Korean CSAT

1. Introduction

The English listening test was included as a part of the English test in

* This work was supported by the Korea Research Foundation Grant funded by the Korean Government (MOEHRD) (KRF-2005-041-101504)

the College Scholastic Ability Test (CSAT) in Korea for the first time in 1994. It was considered an innovative change of the English test in a highly competitive CSAT test towards the measurement of the oral communicative abilities even though the items were not many. The test was introduced in attempt not only to raise the awareness of the importance of students' English oral communicative abilities in English learning but also to lead to the teachers' integration of teaching of listening into English classes and eventually to the change of the English teaching in secondary schools in Korea from exclusively written English to oral English as well by assessing students' listening abilities in the CSAT.

The CSAT listening test has been implemented over 13 years with some changes in the types and the increase of the numbers of the test items. It has been assumed that the test has influenced English teaching and learning because it is a high-stakes test which has important consequences for students. The test score is considered to play an important role in deciding the students' admission to the universities they want to enter in Korea. Therefore, teachers should know well about the contents, types, and format of the test which has been covered in the previous tests and prepare their students to have higher score in the test by attempting to fit their teaching into the test.

However, few studies of the impact of the CSAT test on teaching and learning have been reported up to now. It was only attempted in the research of Kim and O (2002). Nevertheless the study failed to highlight the listening test and seemed to miss the intention of the introduction of a listening part of the test, the central part of the change of the test from previous university entrance tests, which policy makers and test designers believe lead to the change of curriculum and eventually bring about the realization of Communicative Language Teaching in a classroom. Considering the expected role of the listening part of the CSAT test, washback of the listening test should be highlighted and its influences need to be closely examined in more detail by directly looking into secondary school classrooms.

This study was aimed at investigating the washback effects of the Korean CSAT English listening test on English teaching and learning in

a secondary school context. To get to the goal, first of all CSAT English listening test's characteristics were examined. Then the teachers' perceptions about the CSAT listening test and its effects on English teaching and learning were explored. Finally the actual aspects of English teaching and learning practices in secondary English classrooms in association with the test were probed.

2. Literature Review

2.1. The Concept of Washback Effects

Traditionally a test has been used as a tool to measure how well students have achieved the learning objectives at the end of the instruction. More often teachers have utilized tests to examine the degree of their students' achievements towards learning goals which are specified in the curriculum. They design and develop test items or tasks based on the contents and skills they taught in the classes. Given the function of a test as a measurement tool, a test comes at the last stage of the teaching and learning process and is constructed based on the curriculum.

However, some researchers (Alderson, 1986, 1991; Alderson & Wall, 1993; Andrew, 1994, 1995; Cheng, 1997, 2000, 2005; Popham, 1987; Shohamy, 2001) have recently argued that a test can come first before the teaching and learning process and be used as a means to innovate the curriculum. They have noted that tests could be employed as a way to incur some changes of the instruction and lead to top-down educational reform strategies. They have claimed that if some intended change is introduced into the test and the test is high-stakes test which has important consequences for individuals and institutions, teachers as well as students will be very sensitive to the test and interested in the change and the test will have an influence upon teaching and learning.

Alderson (1986) identified such an effect of the test on teaching and learning and referred to it as 'washback' in that the influence would work in a backward direction and cause testing to lead teaching and learning. He also noted the fact that the direction is largely reversed

and asserted that the test could be intended as levers for change and curriculum innovation by influencing the attitudes, behaviors, and motivation of teachers, learners and parents. In a similar vein, Messic (1996) defined the washback as "the extent to which a test influences language teachers and learners to do things they would not necessarily otherwise do that promote or inhibit language learning" (p. 241).

Even though the term 'washback' has been preferably used to refer to the influence of a test on teaching and learning, other similar related terms has often been used together. Biggs (1995) used the term 'backwash' to refer to the same phenomena without clarifying any distinction between the terms. Wall (2005) employed the term 'test impact' which was distinguished from 'washback' in terms of the scope of the effects. According to Wall, 'test impact' was referred to the effects that a test might have on individuals, policies or practices within the classroom or educational system as a whole, whereas 'washback' was defined as the effects of tests on teaching and learning.

The concept of 'washback' was also referred to as 'measurement-driven instruction' by Popham (1987) in that a test can be designed intentionally to bring about change or improvement of curriculum or instruction. In order to achieve this goal, Shepard (1993) emphasized a match between the content and format of the test and those of the curriculum and introduced the term 'curriculum alignment'. In addition, Fredericksen and Collins (1989) used the term 'systematic validity' to describe the situation in which a new or revised examination was introduced into the education system with the aim of improving teaching and learning. Similarly Messick (1996) maintained that a test's validity should be appraised by the degree to which it manifests positive or negative washback and described it 'consequential validity'.

In summary, it has been asserted that tests can be used as a means to control instruction in addition to the traditional evaluative purposes. There has been a conviction that a test has various effects on teaching and learning which is called 'washback'. The concept of washback is defined as any kinds of impact of a test on teaching and learning. The study of washback should go further and explore how much or to what extent a

test has an effect on teaching and learning in a specific educational setting.

2.2. Empirical Studies on Washback Effect

Some empirical studies have been carried out to examine the washback effect of a particular test on teaching and learning within a specific educational context. Alderson and Wall (1993) reviewed the concept of washback and argued that much of the concept had been based on assertion rather than empirical findings. They presented a number of 'Washback Hypothesis' which were meant to illustrate some of the effects that tests might have on teaching and learning. Furthermore they mentioned the methodology of washback study and asserted the importance of data collection not only through surveys or interviews but also through observation on teaching and learning in the classroom.

As an empirical study on washback, Wall and Alderson (1993) examined whether the English test, which was newly introduced test into curriculum in Sri Lanka with the aim of measuring students' English communicative ability, would have effect on teaching and learning. They initially expected that the test would reinforce innovations in teaching materials and curricula and produce positive washback. However, they discovered that the introduction of new test was much more limited than expected. The findings indicated that the test had an impact on teaching content but not on teaching methodology.

Subsequently, Alderson and Hamp-Lyons (1996) did research into the impact of the TOEFL test on teaching. They conducted interviews with teachers and students and observations of two teachers teaching both TOEFL-preparation classes and normal English classes. They found that the test influenced not only teaching content but also teaching methodology. But the results also manifested that teaching methodology was more influenced by an individual teacher's teaching style than by the test itself.

Shohamy, Donitsa-Schmidt and Ferman (1996) studied long-term effects of English test on teaching in Israel. They used questionnaires and structured interviews with teachers, students, and inspectors, as well as analysis of teaching materials. They concluded that washback

could change over time and that the form would depend on several factors: the important of the test, the status of the language, the purpose of the test, the format of the test, and the number of the skills and type of skills tested.

Watanabe (2000) explored the connection between university entrance exam and the prevalence of grammar-translation teaching in Japan. He compared the lessons given by two different teachers to prepare their students for two different university entrance exams. He found out that a test would not affect all teachers in the same way. He considered that the personal characteristics of the teachers such as educational background or beliefs about teaching would play an important role in conducting their lessons.

Cheng (1997, 2000) investigated whether changes to an examination within the context of the Hong Kong secondary school system would actually change different aspects of teaching and learning. The results indicated that teaching content had received the most intensive washback effect. However, the results implied that the data collected were not sufficient enough to provide enough evidence to allow for an analysis of the nature of washback effects, the extent of those effects, and the quality of them.

Qi (2003) reported a study of the intended washback effects of a high-stakes test, the National Matriculation English Test (NMET) in China. Qi compared the test constructors' intentions with the ELT practice in senior students' classroom. It was discovered that there was a mismatch and a partial match between what is intended and what is reported to be the practice. The mismatch concerned the teaching of linguistic knowledge, which showed that the test constructors' intention to decrease the teaching of formal linguistic knowledge was not realized. She concluded that the test produced only limited intended washback effects, as teaching of linguistic knowledge was still emphasized and the kind of language was restricted to the skills tested in the test.

With the empirical studies outside of Korea, Kim and O (2002) studied on the washback of the Korean Scholastic Abilities Test (KSAT), a high-stakes test for state-wide university entrance exam,

and added the empirical evidence to the studies of washback. They examined the relationship between the KSAT English test and characteristics of 12th grade English teaching in Korea. The study revealed the discrepancies between test designers' intentions and what was occurred in the classrooms. However, it was not explained why teachers reacted to the test in the way they did in their classrooms.

Another domestic research on washback is Jung (2006). It investigated washback effects of the nationwide standardized English listening test on English teaching and learning in Korean EFL secondary schools. The notable thing was that the test positively influenced the students' listening abilities. It was found that the listening test had an impact on the students' learning contents and led to the study on the colloquial expressions and the communicative functions. The result implied that the test had the beneficial effect on the change of the curriculum from grammar translation teaching to communicative language teaching as the test designer intended. Despite the positive effect of the test, Jung (2006) argued that the test was not the only factor in incurring the change of learning contents and suggested the need of study on the other factors which influenced the teaching and learning of English as well.

Even though some recent studies of washback were conducted empirically as reviewed above, it was criticized that these were like simple case studies for certain tests (Viphavee, 2001). As stated in Wall (2005), washback is not a simple monolithic phenomena but a multi-dimensional complex one which will be affected by many factors. To induce general theory about washback effect, more factors should be explained by making use of analysis of materials and documents, classroom observations, and individual and group interviews (Cheng & Watanabe, 2004). Furthermore, more research about the same or different tests should be done to prove whether the results are consistent or contradictory with the previous washback studies.

3. Method

3.1. Research questions

This study was aimed at investigating the washback effects of the Korean CSAT English listening test on English teaching and learning in a secondary school context. More specifically the following research questions were posed for the purpose of the study.

- 1) What are the CSAT English listening test's characteristics?
- 2) What are the teachers' perceptions about the listening test and its effects on English teaching?
- 3) What are the actual aspects of English teaching and learning practices in secondary English classrooms in association with the test?

3.2. Subjects

The questionnaires were carried out to 385 high school English teachers mainly teaching in the area of Gwangju Metropolitan City and Jeonnam Province. They were requested to answer the items on the questionnaire. The following table 1 indicates the information for the respondents.

Table 1: Information for the respondents (n=385)

Location	Big city	Small city	Small town	Others
Number (%)	87 (22.6%)	98 (25.4%)	131 (34.3%)	69 (18.0%)
Teaching experience	Less than 5 years	Less than 10 years	Less than 15 years	More than 15 years
Number (%)	76 (19.7%)	127 (33.0%)	118 (30.6%)	35 (9.1%)

As shown in table 1, the teachers were classified in terms of the location where their schools were situated, and the teaching experience they had at schools. With regard to the area of their work, 87 teachers (22.6%) were working at schools in Gwangju Metropolitan city, 98 teachers (25.4%) in cities, 131 teachers (34.3%) in towns, 69 teachers (18.0%) in some isolated areas or islands in the area of Jeonnam Province. Their teaching experiences ranged from less than 5 years to more than 15 years. Those who had less than 15 years and more than 15 years of teaching experiences, the group who was teaching before the CSAT listening test was introduced, numbered 118 (30.6%) and 35 (9.1%), respectively.

3.3. Data Collection

To answer the research question 1, the test items of CSAT English listening tests conducted over 13 years from 1994 to 2006 were collected and analyzed in terms of 5 categories such as language functions, topic, situation, text type, and factual/inferential understanding. For the research question 2, two different instruments for survey such as questionnaire and interview were employed. Questionnaires consisted of 5 subparts such as basic information about respondents, overall perceptions of the test, its effects on teaching contents, teaching activities, and teaching materials. To supplement the questionnaire, interviews were administered to 38 teachers. For the research question 3, the data were obtained through a total of 8 class observations which were randomly selected and observed for 2 different teachers' classes at 4 different schools.

3.4. Data Analysis

To examine the characteristics of the CSAT English listening test, test items from 1994 to 2006 were analyzed in terms of 5 categories such as language functions, topics, situations, text type, factual/inferential understanding. They were itemized and described under each category in a table to show the changes as the years go by. Also the characteristics of the listening test were compared with objectives of listening skills of national curriculum.

Quantitative data from questionnaires were processed with the SPSS (version 12.0) for descriptive statistics and frequency analysis. For the subparts of teachers' perceptions about the test and its effects on listening contents of the questionnaire, 8 items and 5 items were given respectively and answered on a 5 point-likert scale, ranging from 1 to 5 (1= strongly disagree, 5=strongly agree) depending on their degree of agreement. Mean score for each item was calculated and compared among items in the same subpart. On the other hand, for the subparts of teaching activities and materials, teachers' preference or actual use were counted and represented by frequency and percentage.

Qualitative data from interviews and observation were analyzed

following TESOL Quarterly Qualitative Guidelines (TESOL, 2000b). Interview questions were focused on teachers' view regarding positive effects and negative effects of the test. Their answers were analyzed and categorized. Classroom observation journal written in terms of teaching skills, practice time, materials and activities were also analyzed under each category.

4. Results

4.1. Characteristics of CSAT English Listening Test

When CSAT English listening test was introduced in 1994 as a way of promoting Communicative Language Teaching (CLT) in Korean secondary classrooms, it seemed to be attempted very cautiously since a big change in the high-stakes test on short notice could impose a big burden on students as well as teachers. The number of the listening test items comprised only 8 in the tests of 1994 and 1995 and it increased to 10 in the test of 1996 and finally the portion of the listening test increased almost double to 17 items (34%) from the test of 1997 up to now as of the test for 2006 CSAT.

With the increase of the number of the test items, CSAT listening test has been featuring more varied communicative functions and topics since 1997. Table 2, the result of the item analysis of 2000 CSAT English listening test, indicated that the test appeared to have a tendency to be more communicative by including authentic texts such as recorded message, advertisement, announcement, and lecture. In addition, the test was developed to integrate a speaking skill into a listening though it was tested indirectly through multiple-choice technique within the limit of a paper test. And there appeared 4 or 5 test items at the end which ask students to listen to some conversations or passages and choose the most appropriate dialogue which would come next. Moreover the number of items which would ask students' inferential understanding increased and students were supposed to employ a high-level thinking skill to infer intention, situation, reason, purpose, main idea, or relationship from the context while they were listening.

Table 2: Item analysis of 2000 CSAT listening test

Item	Function	Topic	Situation	Text Type	Factual/ Inferential (Question)
1	Describing people	Photo	Classroom	Conversation /Picture	Factual (people)
2	Meeting people	Flight schedule	Airport	Conversation	Inferential (place)
3	Expressing feeling	Feeling		Conversation	Inferential (feeling)
4	Understanding message	Message (favor)	Answering machine	Passage (recorded message)	Factual (calling purpose)
5	Shopping	Money	Sports shop	Conversation	Inferential (money)
6	Suggesting	Transportation	Train station	Conversation	Factual (Transport)
7	Talking about family	Family		Conversation	Factual (reason)
8	Understanding topic	Sleeping Exercise	Lecture	Passage (lecture)	Inferential (topic)
9	Understanding situation	Appreciation	Show	Passage (announcement)	Inferential (relation)
10	Understanding purpose	Train schedule	Train station	Conversation	Factual (purpose)
11	Comparing	Job	Advertise	Conversation (advertisement)	Factual (reason)
12	Understanding reason	TV shows		Conversation	Factual (reason)
13	Matching the picture with a dialogue	Situation	Picture description	Conversation (pictures)	Inferential (situation)
14	Checking out	Library use	Library	Conversation	Inferential (next thing)
15	Asking information	Work	Street	Conversation	Inferential (next thing)
16	Finding direction	Direction	Street	Conversation	Inferential (next thing)
17	Speaking after listening			Passage (situation)	Inferential (next thing)

4.2. Teachers' Perceptions toward the Listening Test

4.2.1. Questionnaire

As shown in Table 3, as to the items of 1, 2, 3, and 4 pertaining to the test's contribution to students' learning of listening, teachers reported that the existence of the listening test in CSAT incurred students' motivation in listening (4.25) and also positively affected student's learning of listening (4.18) and listening ability (3.85) and led to the development of students' communicative ability (3.47). However, as to the items such as 5, 6, 7, and 8 about the test's validity and reliability, they showed a mediocre degree of satisfaction. Their levels of agreement were for reliability (3.26), contents (3.15), item type (3.07), and current test (3.02).

Table 3: Teachers' perceptions or attitudes toward the test

Items	Mean	SD
1. The test helps students develop communicative ability.	3.47	.53
2. The test helps students develop listening ability.	3.85	.65
3. The test promotes students' learning of listening.	4.18	.87
4. The test enhances students' interests or need in listening.	4.25	.72
5. Contents of the test items are valid.	3.15	.74
6. Types of the test items are valid.	3.07	.65
7. The tests are reliable.	3.26	.68
8. The test should be continued as it is.	3.02	.64

(N=385, 1=Strongly disagree, 5=Strongly agree)

As to the items concerning the listening test's effects on teaching contents, Table 4 suggested that the teachers were teaching vocabulary or expressions (4.28), topics (3.85), communicative functions (3.63), pronunciation (3.08), and grammar (2.64) to prepare students for the listening test. This result implied that the teachers were conscious of the contents of the previous listening tests and they were teaching what was mostly covered in the tests. In addition, it reflected that the

teachers believed teaching of sentence structure or grammar was in nearly no relation with the listening test.

Table 4: Teachers' teaching content for listening

Items	Mean	SD
1. Vocabulary or expressions	4.28	.68
2. Topics	3.85	.74
3. Communicative functions	3.63	.69
4. Sentence structure or grammar	2.64	.47
5. Pronunciation or phonological rules	3.08	.53

(N=385, 1=Strongly disagree, 5=Strongly agree)

With regard to teaching activities the teachers used for the test, the teachers' responses were summarized as in Table 5. It was found that the activities teachers preferred were mostly 'Listen and choose the right answer (72.21%)' or 'Listen and fill in the blank (53.25%)'. The reason seemed to be strongly related to the types of test items CSAT English test employed. More importantly, the result suggested that the teachers should use more variety of teaching activities for the listening classes.

Table 5: Teaching activities for listening (N=385)

Items	Frequency	%
1. Listen and fill in the blank	205	53.25
2. Listen and choose the right answer	278	72.21
3. Listen and say the response	94	24.42
4. Listen and match the picture	109	28.31
5. Listen and complete the chart or table	49	12.73
6. Listen and follow the direction	53	13.77
7. Listen and repeat	32	8.31
8. Listen and do a role play	30	7.79

Table 6 displayed teachers' responses regarding teaching materials for listening. Most of the teachers tend to use practice-test books (92.73%) rather than textbooks (30.39%). Pop songs or TV dramas were likely to be used much less and represented only 4.16%, 2.60% respectively. The

result implied that teaching of listening took place only for the test preparation, heavily relying on practice-test books.

Table 6: Teaching materials for listening (N=385)

Items	Frequency	%
1. School textbook	117	30.39
2. Practice-test book or Mock test	357	92.73
3. Pop songs	16	4.16
4. TV drama	10	2.60

4.2.2. interview

The interviews proceeded to identify the teachers' views about the test effects and a set of pre-prepared guided questions were used as an open-ended format. All the responses were summarized as in Table 7 and 8, respectively.

Table 7: Some positive effects of the test (N=38)

Summary of the responses	Frequency	%
1. Awareness of listening skills	31	81.58
2. Development of listening materials	26	68.42
3. Increase of motivation in listening skills	25	65.79
4. Guidelines of learning and teaching of listening	16	42.11

Table 8: Some negative effects of the test (N=38)

Summary of the Answers	Frequency	%
1. Prevents teachers from covering all 4 skills because the test is only for reading and listening skills.	23	60.53
2. Encourage drills for listening but not for genuine listening ability.	15	39.47
3. Familiarize only to the test format	13	34.21
4. Get students to utilize only guessing strategy	8	21.05

The interview results mentioned as positive effects in Table 7 are

compatible with the questionnaire results as in Table 3. It was confirmed that teachers believed the test to have beneficial effects on students' learning of listening, in particular, in awakening them to realize the importance of listening which, otherwise, might have been neglected. Besides, teachers mentioned that they had an idea of what to teach and how to teach for listening by examining the test items or formats to prepare their students for the test.

From the responses in Table 8, it was recognized that the most negative effect of the test is that the test could have teachers and students ignore the skills or contents which would not be on the test. In other words, teachers were concerned that they tended to teach test-taking skills through multiple-choice practice tests to boost the scores but unlikely to promote listening ability.

4.3. Teaching Aspects in the Secondary English Classrooms

To find out the actual aspects of teaching of listening in secondary English classrooms, classroom observation was carried out. Each one of schools in 4 different areas (metropolitan city, city, county, town) was randomly selected. For every school, one class was chosen from regular English classes and the other one was singled out from extra classes. A total of 8 classroom observations were carried out.

Table 9 exhibited the classroom observation journal written in terms of language skills, practice time, materials and activities being used in the classes. With regard to language skills being taught in the regular English classes (Class 1 of each school in Table 9), it was described that reading was the main stream and most of the class time was spent for reading. On the contrary, it was notable that listening was taught briefly during the introduction stage of each unit sometimes together with speaking, while writing skill was rarely covered.

Nevertheless, observation of English classes outside of regular English classes (Class 2 of each school in Table 9) indicated that listening could not be neglected since it is tested in CSAT. Students were forced to

study listening, utilizing morning self-study time before regular classes began or taking away recess time after lunch. Mostly about 20 or 30 minutes were allocated for listening practice almost every day through a school broadcasting system. As for the listening material, EBS listening test preparation books with accompanying CD with them were the most popular, since they were made imitating CSAT listening test format and students could be drilled on the same type of test items as the CSAT listening test.

Table 9: Classroom observation journal

School	Class	Language skills	Practice Time	Materials	Activities	Others
1	Class1	reading, listening with speaking	30min. 10min. beginning	textbook	translation, vocabulary, listen and response	interactional
	Class2	listening	30min. after lunch	practice test book	self-study	school broadcast
2	Class1	reading	40min.	textbook	translation, vocabulary	conventional
	Class2	listening	30min. morning	EBS test book	self-study	school broadcast
3	Class1	reading	40min.	textbook	translation, vocabulary	conventional
	Class2	listening	20min. morning	EBS test book	self-study	school broadcast
4	Class1	reading	40min.	textbook	translation, vocabulary	conventional
	Class2	listening	20min. morning	EBS test book	self-study	school broadcast

5. Conclusion and Suggestions

This study was designed to investigate the washback effects of the Korean CSAT English listening test on English teaching and learning in a secondary school context. The data were collected through the analysis of test items, as well as questionnaire, interview, and observation. The results showed that the test had effects on curriculum, teaching contents, teaching activities, and teaching materials.

The analysis of the development of CSAT listening tests and the test items indicated that the listening skill was likely to play more important part in CSAT English test and the test gradually tended to include more characteristics as a Communicative Language Test. It was clear that the test items covered more variety of communicative functions, or topics and the situations and the texts were adopted from authentic language life. Interestingly, this characteristics of the CSAT listening test seemed to influence the 7th National Curriculum which was implemented from 1998 in a secondary school. It is notable that the test objectives and contents are fairly compatible with the ones about the listening stipulated in the curriculum. This result leads to the conclusion that the test had effects on the curriculum and in part contributed to promotion of Communicative Language Teaching in a secondary classroom.

In another way, some other empirical evidences proved that the listening test had some impacts on teaching and learning practices. Questionnaire analysis indicated that the teachers perceived the listening test had a positive effect on students' motivation, learning of listening, and listening ability, while they showed a mediocre degree of satisfaction about the test reliability, contents, types. In addition, the interview results revealed the teachers' views that the test had not only the positive washback in that it would provide the guidelines of teaching of listening for the teachers and stimulate motivation in listening for the students, but also it had negative washback admitting that teachers tend to teach for the test but not for the genuine learning. The finding was consistent with the previous studies on positive and negative washback (Alderson & Wall, 1993; Cheng, 2005) Finally, the analysis of the class observation journal manifested that teaching of listening took place not with some activity-based tasks in a regular class by a teacher but with test preparation books, imitation copies of the CSAT listening test, in a self-study class through school broadcast system.

Based on the above conclusions, several suggestions for CSAT English listening test can be made to bring more beneficial effects to

English teaching and learning in secondary school classrooms. First, the types of listening test items need to be varied. Currently the test tends to heavily rely on the type of 'listen and choose'. However, in reality listening is likely to take place accompanying with other responses such as 'listen and act' 'listen and fill in the form' and 'listen and say'. If the test includes different types of test items which are similar to the listening of real life, teaching activities for English listening in a classroom will be varied too.

Secondly, listening test items should be developed incorporating other language skills so that speaking or writing skills can be dealt with together in a classroom. If the test contains integrative test items like 'listen and say' or 'listen and write', teachers will prepare their students to do well on the test items which require overall language skills. It will affect English teaching aspects of secondary schools in which regular class time is spent in mostly reading and other language skills are hardly covered. It is expected that the integrative test will eventually contribute to promoting students' speaking and writing skills in addition to listening.

References

- Alderson, J. C. (1986). Innovations in language testing. In M. Portal (ed.), *Innovations in language testing: Proceedings of the IUS/NFER conference* (pp. 93-105). London: NFER/Nelson.
- Alderson, J. C. (1991). *Language testing in the 1990s: How far have we come? How much further have we to go?* Retrieved June 2, 2003, from ERIC EDRS, ED 365145.
- Alderson, J. C., & Hamp-Lyons, L. (1996). TOEFL preparation courses: A study of washback. *Language Testing, 13*(3), 280-297.
- Alderson, J. C., & Wall, D. (1993). Does washback exist? *Applied Linguistics, 14*(2), 115-129.
- Andrews, S. (1994). The washback effect of examinations: Its impact upon curriculum innovation in English language teaching.

Curriculum Forum, 4(1), 44-58.

- Andrews, S. (1995). Washback or washout? The relationship between examination reform and curriculum innovation. In D. Nunan, V. Berry, & R. Berry (Eds.), *Bringing about change in language education* (pp. 67-81). Hong Kong: Dept of Curriculum Studies, University of Hong Kong.
- Biggs, J. B. (1995). Assumptions underlying new approaches to educational assessment, *Curriculum Forum*, 4(2), 1-22.
- Cheng, L. (1997). How does washback influence teaching?: Implications for Hongkong. *Language and Education*, 11, 38-54.
- Cheng, L. (2000). *Washback or backwash: A review of the impact of testing on teaching and learning*. Retrieved June 2, 2003, from ERIC EDRS, ED 442280.
- Cheng, L. (2005). Changing language teaching through language testing: a washback study. *Studies in Language Testing*, 21. Cambridge: Cambridge University Press.
- Cheng, L., & Watanabe, Y. (eds.). (2004). *Washback in language testing*. New Jersey: Lawrence Erlbaum Associates.
- Frederiksen, J. R., & Collins, A. (1989). *A system approach to educational testing*. *Educational Researcher*, 18(9), 27-32.
- Jung, H. (2006). The washback effects of standardized English listening test conducted by municipal or provincial offices of education. *The Linguistic Association of Korea Journal*, 14(3), 177-194.
- Kim, D. J., & O, K. (2002). *Washback on 12th grade English classes of the English tests within Korean university entrance exams*. *English Teaching*, 57(3), 303-329.
- Messick, S. (1996). Validity and washback in language testing, *Language Testing*, 13, 241-256.
- Popham, W. J. (1987). The merits of measurement-driven instruction. *Phi Delta Kappa*, 68, 679-682.
- Qi, L. (2004). Has a high-stakes test produced the intended changes? In L. Cheng & Y. Watanabe (eds.), *Washback and language testing* (pp. 171-190). New Jersey: Lawrence Erlbaum Associates.
- Shepard. L. (1993). The place of testing reform in educational reform: A

reply to Cizek, *Educational Researcher*, 22(4), 10-14.

- Shohamy, E. (2001). *The Power of tests: A critical perspective on the use of language tests*. Harlow: Pearson Education.
- Shohamy, E., Donitsa-Schmidt, S., & Ferman, I. (1996). Test impact revisited: Washback effect over time, *Language Testing*, 13, 298-317.
- TESOL. (2000b). TESOL Quarterly qualitative guidelines. *TESOL Quarterly* 31(2), 387-388.
- Viphavee, V. (2001). An interview with J. Charles Alderson. *Issues in Applied Linguistics*, 12(1), 91-109.
- Wall, D. (2005). The impact of high-stakes examinations on classroom teaching. *Studies in Language Testing*, 22. Cambridge: Cambridge University Press.
- Wall, D., & Alderson, J. (1993). Examining washback: The Sri Lankan impact study. *Language Testing*, 10, 41-70.
- Watanabe, Y. (2000). Washback effects of the English section of Japanese university entrance examination on instruction in pre-college level EFL, *Language Testing Update* 27, 42-67.

Haeng Jung
Department of English Language and Literature
Honam University
330 Eodeung-ro, Gwangsan-gu
Gwangju, 506-090, Korea
Phone: 82-62-940-5523
Email: hjung58@dreamwiz.com

Received: 30 June, 2008

Revised: 22 August, 2008

Accepted: 7 September, 2008